International Journal of Advanced Engineering Research and Science (IJAERS) [Vol-4, Issue-4, Apr- 2017]
https://dx.doi.org/10.22161/ijaers.4.4.27 ISSN: 2349-6495(P) | 2456-1908(O)

Forecasting Particulate Matter Concentrations: Use of Unorganized Machines

Yara de Souza Tadano[1], Hugo Valadares Siqueira[1], Thiago Antonini Alves[1], Manoel Henrique de Nobrega Marinho[2]

[1]Federal University of Technology – Paraná, Ponta Grossa/PR, Brazil
[2]Polytechnic School of Pernambuco, Pernambuco University, Recife/PE, Brazil

***Abstract*— *Air pollution is an environmental issue studied worldwide, as it has serious impacts on human health. Therefore, forecasting its concentration is of great importance. Then, this study presents an analysis comprising the appliance of Unorganized Machines – Extreme Learning Machines (ELM) and Echo State Networks (ESN) aiming to predict particulate matter with aerodynamic diameter less than 2.5 μm ($PM_{2.5}$) and less than 10 μm (PM_{10}). The databases were from Kallio and Vallilla stations in Helsinki, Finland. The computational results showed that the ELM presented best results to $PM_{2.5}$, while the ESN achieved the best performance to PM_{10}.*

***Keywords*— air pollution, neural networks, particulate matter.**

I. INTRODUCTION

Air pollution has always been an environmental issue worldwide. Therefore, its prediction comprises an important topic, mainly due to its impact on human health. Among urban air pollutants, particulate matter (PM) has been considered one of the most harmful ones, as it is related to hospital admissions for respiratory and cardiovascular problems, and even death [1], [2], and [3]. Then, PM concentration forecasting is of great interest to government plans and to warn population regarding events of severe pollution levels.

Therefore, the present study aims to show a brief comparative analysis of artificial neural networks performance, the well-known Unorganized Machines – Extreme Learning Machines (ELM) and Echo State Network (ESN) – on predicting particulate matter with aerodynamic diameter less than 10 □m (PM_{10}) and particulate matter with aerodynamic diameter less than 2.5 □m ($PM_{2.5}$). The database comprises Helsinki, Finland, air pollution from two distinct stations, Kallio and Vanilla from 2001 to 2003.

II. PREDICTOR MODELS

The Extreme Learning Machines (ELM) and the Echo State Networks (ESN) – collectively known as Unorganized Machines (UMs) – are artificial neural networks architectures characterized by a simple training process allied to good results [4]. The most important characteristic of these networks is its hidden layer stands untrained, allowing them to train only the output layer in a minimum mean square error sense, which confers a very fast adjust process to the networks [5].

The ELMs, proposed by [6], are feedforward networks, quite similar to the traditional Multilayer Perceptron (MLP) [7]. The authors proved by a constructive approach that the output error of a signal always can be reduced with the insertion of a new neuron in the hidden layer of a feedforward network. A condition have to be respected: the activation function of these neurons needs to be differentiable. By means of a rigorous mathematical demonstration, the authors proved that the structure have generalization capability and are universal approximators. Then, to predictions tasks, the ELM may present adequate results even to unknown input data, when it is trained. The most common way to adjust the weights of the output layer is the application of the Moore-Penrose pseudoinverse operation, which guarantees the best solution by means of a deterministic solution [5].

Unlike the ELM, the ESN are recursive networks endowed by feedback loops of information. It means some output responses are reinserted in the network input, generating an intrinsic memory. This characteristic may be good to solve problems in which the samples present temporal dependence. The ESNs were proposed by [8] and, as the ELM, are universal approximators with generalization capability [8] and [9].

The main difficulty in classic recurrent neural networks application is the training process, once it is necessary to apply nonlinear optimization techniques. This procedure may lead to instability, local convergence and, in general, it has high computational cost. As mentioned, in the ESN, the intermediate layer weights – called dynamic reservoir – remains unadjusted. The theoretical element that guarantees the presence of memory is the echo state propriety, which says the most recent historic of inputs

https://dx.doi.org/10.22161/ijaers.4.4.27 ISSN: 2349-6495(P) | 2456-1908(O)

rules the internal dynamic of the reservoir. The immediate consequence is the weights of this layer may be defined previously. Then, the adjustment process can be limited to the output layer training, in the same way of the ELM case [8].

III. METHODOLOGY

The data preprocessing is very important to the direct application of the neural networks. For this, it is applied the padronization described in Equation (1) [10]:

$$z_i = \frac{d_i - \bar{d}}{\sigma}, \quad (1)$$

where, i =1, ..., N is the index of each sample, $\bar{d}$ is the sample mean and σ is the standard deviation. The new series, z, are stationary with zero mean and standard deviation equals to one. The used lags were defined by preliminary tests and the lags selected are 2, 3, 4, 5, and 9, to PM_{10} and, 1, 2, 4, 8, 9, and 10, to $PM_{2.5}$.

The performance metrics adopted were the Mean Square Error (MSE) and the Mean Absolute Percentage Error (MAPE), presented in Equations (2) and (3) [10]:

$$MSE = \frac{1}{N}\sum_{n=1}^{N}(d_n - y_n)^2, \quad (2)$$

$$MAPE = \frac{100}{N}\sum_{n=1}^{N}\left|\frac{d_n - y_n}{d_n}\right|, \quad (3)$$

where, N is the number of samples, y_n the n-th data predicted and d_n the desired response to the respective predicted data.

IV. COMPUTATIONAL RESULTS

The prediction of PM_{10} and $PM_{2.5}$ concentrations were held to Helsinki, Finland. To model and predict PM time series, the most common method comprises a weighted combination of the same variable observed data. The database consists of daily data from 2001 to 2003 to Kallio and Vallila Stations [11]. These stations are located in regions with distinct characteristics. Kallio is an urban background and Vallila is located in city downtown with influence of traffic. It means, population around Kallio station is less exposed to air pollution then Vallila ones, which are severely exposed to it.

The computational results considered the forecasting with horizon of one-step ahead, and they are presented in Table 1, which shows the average of 30 simulations. The metrics used to evaluate the performance of the neural networks are MSE and MAPE to real domain (in the magnitude of the original data), and only MSE to padronized data. The label "NN" is the number of neurons in the intermediate layer of the neural networks that reached the best performance. The "K" label is related to Kallio station and "V" one is to Vallila station.

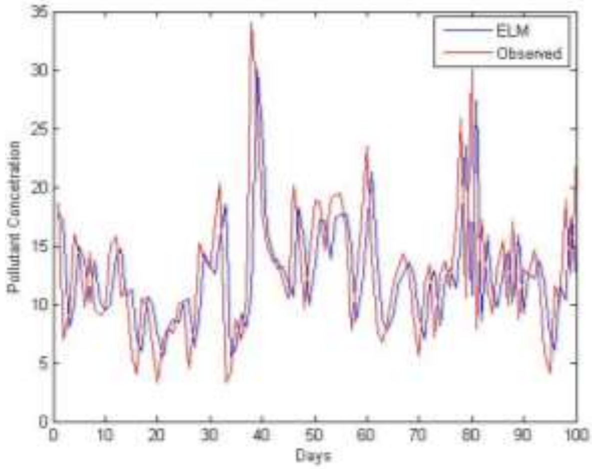
The first observation is that there is no direct relation between performance and number of neurons (processor units). The ESN used always a few number of neurons then the ELM. Interestingly, to PM_{10} forecasting, the ELM achieved the best performance, while to $PM_{2.5}$, the opposite was verified, the ESN showing the best prediction. It may indicate which type of neural network – feedforward or recursive – is more suitable to solve the task. It is important to observe that the Friedman's test [12] was performed to analyze the statistical significance of the results. The p-values found were close to zero, which allows assuming the hypothesis that changes in the predictor lead to distinct results.

Table.1: Mean Square Error (MSE) and Mean Absolute Percentage Error (MAPE) of computational results to PM_{10} and $PM_{2.5}$ using ELM and ESN to Kallio (K) and Vallila (V) stations

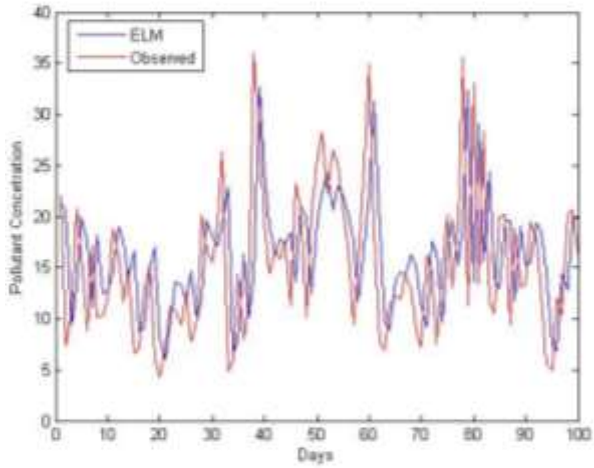
	Metrics	ELM	ESN		ELM	ESN
K- PM_{10}	NN	120	3	K- $PM_{2.5}$	10	3
	MSE padron.	**0.0052**	0.0057		0.0065	**0.0046**
	MSE	**33.1995**	36.5808		21.1057	**14.8165**
	MAPE	**32.5936**	34.7290		42.7161	**37.8863**
V- PM_{10}	NN	70	3	V- $PM_{2.5}$	70	3
	MSE padron.	**0.0029**	0.0034		0.0046	**0.0033**
	MSE	**55.0645**	64.0550		22.6697	**16.4340**
	MAPE	**33.6415**	43.3059		39.9475	**35.7573**

Figure 1 presents the forecasting results in comparison to the observed data to the best cases found. It shows the results to PM_{10} from Kallio and Vallila stations and to $PM_{2.5}$ from both stations, respectively. It is possible to observe that ELM and ESN fitted well to the observed data.

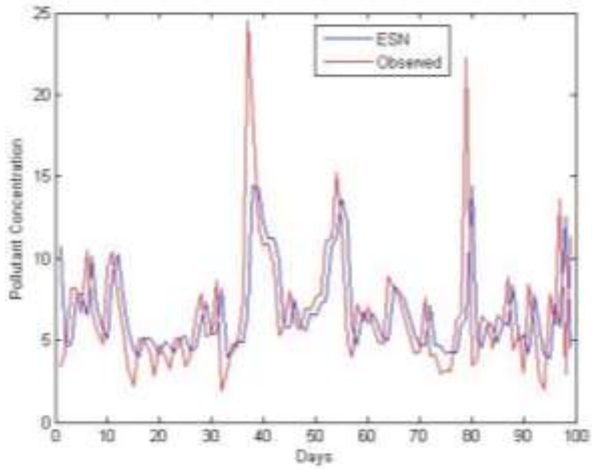
https://dx.doi.org/10.22161/ijaers.4.4.27 ISSN: 2349-6495(P) | 2456-1908(O)



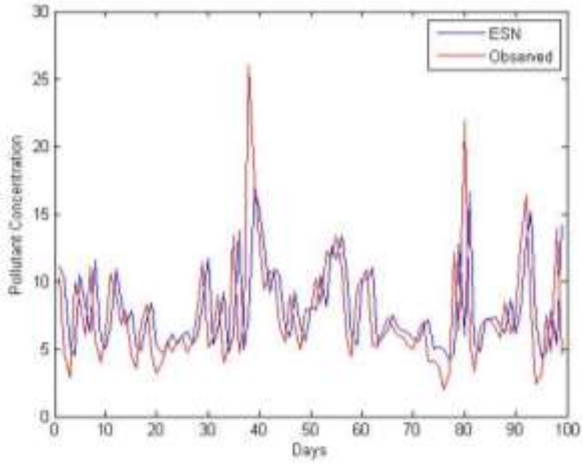
(a) *Kallio and PM_{10}*



(b) *Vallila and PM_{10}*



(c) *Kallio and $PM_{2.5}$*



(c) *Vallila and $PM_{2.5}$*

Fig. 1: Best predictions achieved by the unorganized machines in [□g/m^3]

V. CONCLUSIONS

The main reason to use ELM and ESN architectures is the good results in forecasting problems described in the literature, allied to a simple and efficient training process. The computational results showed that the ELM achieved the best performances to predict PM_{10} concentrations, while the ESN showed better results to $PM_{2.5}$. These results could assist government with prediction tools that could help with mitigating measures. New researches can be done using regularized ELM and ESN with nonlinear output layers and, even applying variable selection techniques can increase the general performance of the models.

REFERENCES

[1] J.P. Langrish, X. Li, S.F. Wang, M.M.V. Lee, G.D. Barnes, M.R. Miller, F.R. Cassee, N.A. Boon, K. Donaldson, J. Li, L. Li, N.L. Mills, D.E. Newby, and L.X. Jiang, "Reducing personal exposure to particulate air pollution improves cardiovascular health in patients with coronary heart disease," Environ. Health Perspect., vol. 120, pp. 367-372, March 2012.

[2] C.F. Wu, Y.R. Li, I.C. Kuo, S.C. Hsu, L.Y. Lin, and T.C. Su, "Investigating the association of cardiovascular effects with personal exposure to particle components and sources," Sci. Total Environ., vol. 431, pp. 176-182, August 2012.

[3] Y.S. Tadano, H.V. Siqueira, and T. Antonini Alves, "Unorganized machines to predict hospital admissions for respiratory diseases" in Proceedings of Latin American Conference on Computational Intelligence (2016).

[4] L. Boccato, E.S. Soares, M.M.L.P. Fernandes, D.C. Soriano, and R. Attux, "Unorganized machines: from Turing's ideas to modern connectionist approaches," Intern. J. of Nat. Comp. Research, vol. 2, pp. 1-16, October 2011.

[5] H.V. Siqueira, L. Boccato, R. Attux, and C. Lyra, "Unorganized machines for seasonal streamflow series forecasting," Int. J. of Neural Systems, vol. 24, pp. 1430009.1-1430009.16, February 2014.

[6] G.-B. Huang, G.-Y. Zhu, and C.-K. Siew, "Extreme Learning Machine: theory and applications," Neurocomputing, vol. 70, pp. 489-501, December 2006.

[7] S. Haykin, Neural Networks and Learning Machines, 3rd ed., New York: Prentice-Hall, 2008.

[8] H. Jaeger, The Echo State Approach to Analyzing and Training Recurrent Neural Networks, Bremem: German Nat. Res. Center for Inf. Tech., Tech. Rep. GMD Report 148, 2001.

[9] A.M. Schäfer, and H.G. Zimmermann, "Recurrent neural networks are universal approximators," Int. J. of Neural Systems, vol. 17, pp. 253-263, August 2007.

[10] G.E.P. Box, G.M. Jenkins, G.C. Reinsel, and G.M. Ljung, Time Series Analysis: Forecasting and Control, 5th ed., New York: John Wiley & Sons, 2015.

[11] D. Voukantsis, K. Karatzas, J. Kukkone, T. Räsänen, A.Karppinen, and M.Kolehmainen, "Intercomparison of air quality data using principal component analysis, and forecasting of PM_{10} and $PM_{2.5}$ concentrations using artificial neural networks, in Thessaloniki and Helsinki," Sci. Total Envrion, vol. 409, pp. 1266-1276, March 2011.

[12] I.R. Luna, and R. Ballini, "Top-down strategies based on adaptive fuzzy rule-based systems for daily time series forecasting," Int. J. of Forecasting, vol. 27, pp. 1-17, July 2011.